

AN ABSTRACT OF THE THESIS OF

Jing Sun for the degree of Honors Baccalaureate of Science in Microbiology presented on May 29, 2008. Title: Transcription Factors Associated with Tomato Seed Germination.

Abstract

approved: _____

Hiroyuki Nonogaki, Associate Professor

Seed germination is strictly defined as the physiological events before radicle protrusion through the endosperm cap, the endosperm region adjacent to the radicle tip. There are two opposing forces governing seed germination: the growth potential of the embryo and the mechanical resistance of the endosperm cap. The endosperm cap cell wall is composed of galactomannan, which is degraded mainly by endo- β -mannanase. *LeMAN2* encodes an endo- β -mannanase which is induced by gibberellin (GA) during germination. The upstream regulators of *LeMAN2* are unknown in tomato. Using the information available for GA-inducible transcription factors in *Arabidopsis*, tomato orthologues of GA-inducible transcription factors were identified and characterized. Four transcription factors are found to be expressed in germinating seeds. One of these genes, *GATA2* exhibited stage- and tissue-specific expression similar to that of *LeMAN2*. *GATA2* contains a GATA zinc finger domain similar to the zinc finger domain of AGP1 in tobacco which binds the AG motif in the promoter region of *NtMYB2*, a Myb protein. An orthologue of *NtMYB2* is expressed in tomato seed during germination. The tomato *GATA2* is hypothesized to be involved in the induction of *LeMAN2* in an indirect manner through the regulation of a Myb protein.

Key Words: endo- β -mannanase, GATA2 zinc finger protein, Myb protein, tomato seed germination, transcription factor.

Corresponding e-mail address: sunj@onid.orst.edu

© Copyright by Jing Sun

May 29, 2008

All Rights Reserved

Transcription Factors Associated with Tomato Seed Germination

by

Jing Sun

A THESIS

Submitted to

Oregon State University

University Honors College

in partial fulfillment of
the requirements for the
degree of

Honors Baccalaureate of Science in Microbiology (Honors Scholar)

Presented May 29, 2008
Commencement June 2008

Honors Baccalaureate of Science in Microbiology project of Jing Sun presented on May 29, 2008.

APPROVED:

Hiro Nonogaki, representing Horticulture

Linda Bruslind, representing Microbiology

Ruth Martin, representing USDA-ARS

Dean, University Honors College

I understand that my project will become part of the permanent collection of Oregon State University, University Honors College. My signature below authorizes release of my project to any reader upon request.

Jing Sun, Author

ACKNOWLEDGEMENTS

I am grateful to the Howard Hughes Medical Institute for the Summer Undergraduate Research grant and the Ernest and Pauline Jaworski Fund for Summer Research Experiences for Underserved Undergraduates in Plant Science grant. I would like to thank Drs. Kevin Ahern and Christopher Mathews for organizing the HHMI program at Oregon State University and their time and advice. I am grateful for Dr. Kent J. Bradford, University of California, Davis for providing tomato seeds. I would like to thank my colleagues Dr. Po-Pu (Eric) Liu, Jessica R. Hewitt, M.Sc., Tanja Homrichhausen, Dr. Masa Asahina, Dr. Wioletta Pluskota, Tammy Chan, and Jennifer Coppersmith for their advice, suggestions, discussions and friendship. I would like to express my sincerest appreciate to my mentor, Dr. Hiro Nonogaki for his tireless patience, guidance and support. I am grateful to my committee members, Drs. Linda Bruslind and Ruth Martin for their time and advice.

CONTRIBUTION OF AUTHORS

Drs. Masa Asahina, Wioletta Pluskota, Department of Horticulture at Oregon State University, and Tammy Chan assisted in the experiments and the completion of this thesis research.

TABLE OF CONTENTS

	<u>Page</u>
Chapter 1. General Introduction.....	2
Chapter 2. Transcription Factors in Tomato Seeds.....	6
INTRODUCTION.....	7
MATERIALS AND METHODS.....	8
<i>Plant materials</i>	8
<i>Tissue print</i>	8
<i>Real-time PCR</i>	9
<i>Bioinformatic analysis</i>	9
<i>RNA probe synthesis</i>	9
<i>Reverse transcription (RT) PCR</i>	9
RESULTS.....	10
DISCUSSION.....	16
Chapter 3. General Conclusion.....	19
Bibliography.....	20

LIST OF FIGURES

<u>Figure</u>	<u>Page</u>
1. The morphology of seed	3
2. Schematic representation of the opposing forces of seed germination.....	4
3. Stage-specific expression of the four transcription factors in tomato seeds during germination.....	11
4. Photographs and schematic representation of the four different tomato seed parts used for RNA extraction and gene expression analysis.....	11
5. Tissue-specific expression of the four transcription factors in tomato seeds during germination.....	12
6. Tissue printing of <i>GATA2</i> expression.....	12
7. Alignment of the amino acid sequences of the tomato <i>GATA2</i> (<i>SlAGP1</i>) and tobacco <i>AGP1</i> (<i>NtAGP1</i>).....	13
8. Alignment of the nucleotide sequences of the tomato <i>Myb2</i> (<i>SlMyb2</i>) and tobacco <i>NtMyb2</i>	14
9. <i>SIMYB2</i> expression in tomato seeds.....	16
10. The hypothetical schematics of <i>LeMAN2</i> regulation by the transcription factor <i>GATA2</i>	17

DEDICATION

I would like to dedicate this work to my loving family, my mother Amy Ai-Ling Zhang and my brother William Sun. Thank you for all your support, encouragement, and unconditional love throughout all these years. I would like to dedicate this manuscript to Hiro Nonogaki for all his patience and desire to teach me not only about research and seed biology but also about life in general. Thank you for your tireless work to help me prepare for my future.

Transcription Factors Associated with Tomato Seed Germination

Chapter 1.

General Introduction

Jing Sun

Seeds maintain vital functions upon which our daily lives are dependent. Seeds are the basis for agriculture, which is an important factor in the world economy. Crops such as maize, soybean, wheat and rice are the major dietary staples in nations around the world and therefore humans have become dependent on seeds as a direct and indirect food source (Grant et al., 1983). Seeds are the primary dispersal unit of plants which enables successful propagation and maintenance of plant species. As the source of nutrition and protection of the embryo, seeds allow the establishment of the next generation of plants. Plants are an important component of the environment and ecosystem around us. Ecosystems are complex webs of animal and plant life. Plant propagation through seeds is essential in maintaining this delicate balance in nature. Seeds are the delivery system of genetic information for the adult plant. Through the genetic information carried by seeds, the adult plants exhibit various phenotypes. For example, crop characteristics such as drought tolerance, hardiness and fruit size are determined by the genetic information contained in seeds.

As the delivery system of genetic information and the dispersal unit for future progeny, seeds have evolved specific morphology and structure designed to protect the embryo and provide the highest probability for survival (Fig. 1).

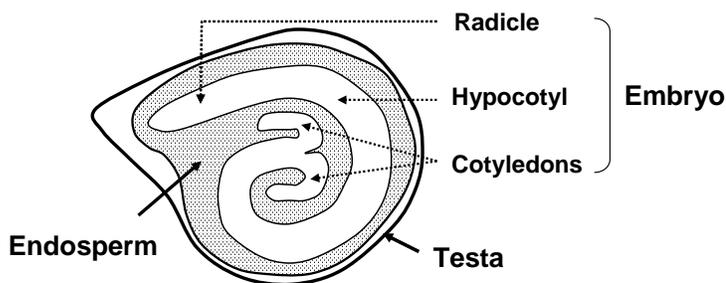


Figure 1. The morphology of seeds. Example of tomato seed is shown. Seed is designed to ensure the proper protection for the embryo. The embryo is surrounded by the endosperm and testa (seed coat).

Seeds are composed of three main tissues: the embryo, endosperm, and testa. The embryo is derived from a zygote, a single cell after fertilization, and will develop into the next generation of plants. The embryo consists of the radicle, hypocotyl, and cotyledons. The radicle protrudes through the endosperm and testa and develops into the main root. The hypocotyl functions in the transfer of nutrients to and from the cotyledons. Cotyledon is an embryonic leaf and can be one (monocotyledonous plants) or two (dicotyledonous plants). The endosperm functions as a nutrient reserve for the developing embryo. The endosperm also provides mechanical resistance to prevent radicle emergence (Watkins et al., 1985; Groot and Karssen, 1987; Sanchez et al., 1990; Bewley, 1997). The testa, the outermost tissue of a seed which is also known as the seed coat, provides protection to the embryo during seed dispersal, storage and germination. The testa also functions in seed dormancy and longevity (Debeaujon and Koornneef, 2000; Debeaujon et al., 2007).

As the beginning for most plants, seed germination is a critical step in the plant life cycle. Seed germination is strictly defined as the physiological events prior to radicle protrusion through the endosperm and testa (Nonogaki et al., 2007). Seeds must first break the desiccated state by imbibition. After water uptake, seeds begin germination unless they are dormant. Germination is completed with radicle emergence. Germination is characterized by two opposing forces: the mechanical resistance of the endosperm and testa and the growth potential of the embryo (Fig. 2).

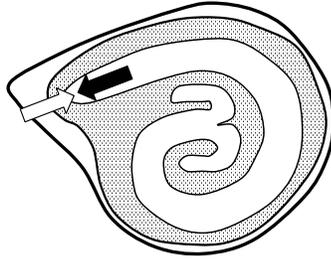


Figure 2. Schematic representation of the opposing forces of seed germination. The embryo growth potential (closed arrow) must be greater than the mechanical resistance of the endosperm and testa (open arrow) in order for radicle to protrude.

If the mechanical resistance of the endosperm decreases, the growth potential increases, or both occur simultaneously, the radicle tip is able to protrude through the endosperm and testa and the seed completes germination. In any scenario, the growth potential of the embryo must exceed the mechanical resistance of the endosperm in order for the seed to germinate.

While the forces behind the mechanisms of seed germination have been understood, the molecular mechanisms underlying the regulation of seed germination are not yet clear. Regulatory proteins such as transcription factors need to be characterized. This thesis focused on the investigation of transcription factors involved in seed germination. Transcription factors and their roles in germination were studied using tomato seed which has become a model system for seed germination research. The entire tomato genome has not yet been sequenced. Therefore, translation of the information obtained from the model plant *Arabidopsis* to tomato was attempted. Thus, this research focused on translational biology which has potential to aid in agriculture and food production in the future.

Chapter 2

Transcription Factors in Tomato Seeds

Jing Sun

INTRODUCTION

Seed germination is completed by radicle emergence which is a consequence of an increase in embryo growth potential: the embryo overcomes the mechanical resistance of the endosperm. In addition to the increase in the embryo growth potential, a weakening of the endosperm is necessary for germination to occur (Cantliffe et al., 1984; Groot and Karssen, 1987). In the micropylar region of the endosperm (endosperm cap), the cell wall provides mechanical resistance to the radicle (Groot and Karssen, 1987; Chen and Bradford, 2000). The cell wall of tomato endosperm cap is composed of mannan polysaccharides, galacto-, gluco-, or galactoglucomannans (Groot et al., 1988). The cleavage of these mannan chains weakens the cell wall and allows radicle protrusion. A few enzymes are known to modify the galactomannan chains: α -galactosidase, β -mannosidase, and endo- β -mannanase. The enzyme α -galactosidase cleaves the galactoses from the main mannan backbone, β -mannosidase removes the mannose residues from the mannan chain terminus and endo- β -mannanase cleaves the mannan chain internally (Bewley and Black, 1994).

The gene *LeMAN2* (*Lycopersicon esculentum* [currently *Solanum lycopersicum*] mannanase) is known to encode an endo- β -mannanase. Tissue printing of wild-type tomato seeds and hybridization with the *LeMAN2* probe have shown that *LeMAN2* is expressed primarily in the endosperm cap (Nonogaki et al., 2000). It is known that *LeMAN2* is upregulated by the plant hormone gibberellin (GA). GA plays important roles in many biological processes including seed germination (Crozier et al., 2000; Yamaguchi, 2008). Molecular mechanisms of *LeMAN2* upregulation and the induction

of seed germination by GA are unknown in tomato seed. Upstream regulators of *LeMAN2* in tomato seeds have not been identified.

In the model plant *Arabidopsis thaliana*, transcription factors induced by GA during seed germination have been identified (Ogawa et al., 2003). These transcription factors are hypothesized to control downstream genes associated with seed germination. Orthologues of these *Arabidopsis* transcription factors in tomato are potentially involved in the regulation of seed germination. Bioinformatic analysis using databases such as the Expression Sequence Tag (EST) Databases was performed to identify the tomato orthologues corresponding to these *Arabidopsis* transcription factors. In this chapter, the characterization of tomato transcription factors is described.

MATERIALS AND METHODS

Plant materials

Tomato seeds (*Solanum lycopersicum* cv. Moneymaker) were imbibed in water in 9 cm plastic Petri dishes with two layers of filter paper (No. 2 Whatman Inc., Clifton, NJ). Tomato seeds were incubated at 25°C for 12-48 hours (h).

Tissue print

Tomato seeds were bisected with a razor blade and the seed halves were pressed onto a positively charged membrane for about 15 seconds (s) (Nonogaki and Bradford, 2003). The tissues were removed and the membrane was UV cross-linked. The membrane was pre-hybridized for 15 minutes at 60°C in hybridization buffer (Nonogaki et al., 2000; Nonogaki and Bradford, 2003) before being hybridized for 16-18 h at 60°C

in the hybridization buffer with 100 ng mL⁻¹ Digoxigenin (DIG) labeled RNA probes. The membrane was blocked with 5% nonfat milk solution and washed. Alkaline phosphatase-conjugated anti-DIG antibody was used to detect the RNA probes colorimetrically.

Real-time PCR

First-strand cDNA was synthesized from 1 µg of total RNA with QuantiTect Reverse Transcription Kit according to the manufacturer's instructions (Qiagen, Valencia, CA). Quantitative reverse transcription PCR with Taq-Man technology was performed using first-strand cDNA as a template on a sequence detector system (ABI PRISM 7000; Applied Biosystems, Foster, CA). Results were normalized using an 18S rRNA as the internal control.

Bioinformatic analysis

Tomato sequences were searched via unigenes in the EST databases (<http://www.sgn.cornell.edu/> and <http://compbio.dfci.harvard.edu/tgi/cgi-bin/tgi/gimain.pl?gudb=tomato>). The sequence of *NtMyb2* in tobacco (*Nicotiana tabacum*) was obtained from NCBI nucleotide database (<http://www.ncbi.nlm.nih.gov>). Sequence alignment was performed using NCBI BLAST searches and the Oregon State University Center for Genome Research and Biocomputing (CGRB) alignment programs.

RNA probe synthesis

DIG-labeled RNA probe was synthesized using MAXIscript T7/T3 Kit (Ambion, Austin, TX) using 1 µg of *Pst*I digested DNA and T7 RNA polymerase. RNA synthesis was performed at 37 °C for 2 h.

Reverse transcription (RT) PCR

Total RNA was extracted from 24 h imbibed tomato seeds using standard phenol-SDS extraction (Sambrook et al., 1989) and 2 µg of DNase-treated total RNA was used for reverse transcription with a RETROscript kit (Ambion, Austin TX). The RT product was amplified by PCR using *Ex Taq* DNA polymerase (Takara, Madison, WI) and primers for the tomato Myb gene (*SIMyb2*, see RESULTS) (5' GTTAGAGCTCCTTGTTGTGA 3' and 5' CACATCACCATCCACTATAC 3'). The following conditions were used for PCR: the initial denaturation at 94 °C (4 min), 25 cycles at 94 °C (1 min), 50 °C (1 min) and 72 °C (2 min) followed by extension at 72 °C (7 min).

RESULTS

Tomato transcription factors that are potentially involved in seed germination control were selected based on the information of the Arabidopsis seed germination-associated transcription factors (Ogawa et al., 2003) and tomato unigene sequences (<http://www.sgn.cornell.edu/>). Four transcription factors were confirmed to be expressed

during germination of tomato seeds by real time polymerase chain reaction (real time PCR) (data not shown).

Stage-specific expression of the four transcription factors was analyzed using quantitative real time PCR. While three transcription factors, *GATA1* (*GATA zinc finger protein 1, C2C2 type*), *MBF* (*Multiprotein Bridging Factor*) and *AS2* (*Asymmetric 2*) were expressed at relatively early stages (12 h imbibition), one gene *GATA2* (*GATA zinc finger protein 2, C2C2 type*) showed highest expression at the relatively later stage (36 h) which was similar to the timing of *LeMAN2* expression (Fig. 3).

To examine tissue-specificity of the expression of the four transcription factors, RNA was extracted from four different parts of tomato seeds: endosperm cap (EC), lateral endosperm (LC), cotyledon-half embryo (C), and radicle- half embryo (R) (Fig.4). *GATA2* was expressed primarily in the endosperm, although low level expression of this gene was also detected in radicle tissue (Fig. 5). The tissue localization of *GATA2* was confirmed by tissue printing (Fig. 6).

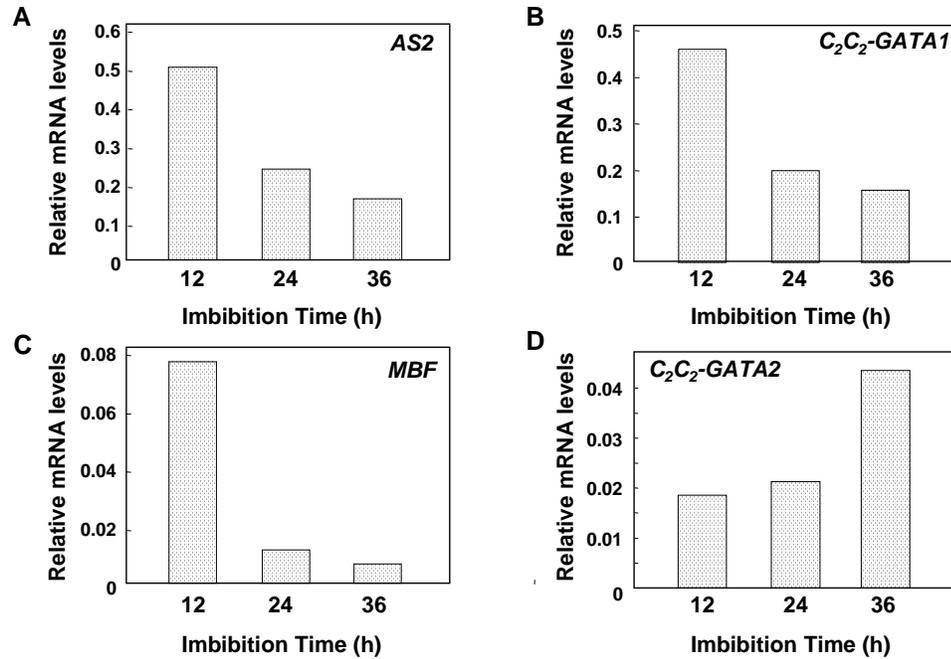


Figure 3. Stage-specific expression of the four transcription factors in tomato seeds during germination. A, B, C and D, the expression of *GATA1*, *GATA2*, *MBF* and *AS2*, respectively. Expression was analyzed by quantitative real time PCR.

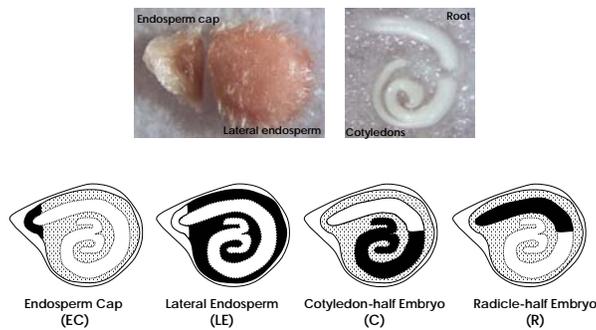


Figure 4. Photographs and schematic representation of the four different tomato seed parts used for RNA extraction and gene expression analysis. EC: endosperm cap, LE: lateral endosperm, C: cotyledon-half embryo, R: radicle-half embryo.

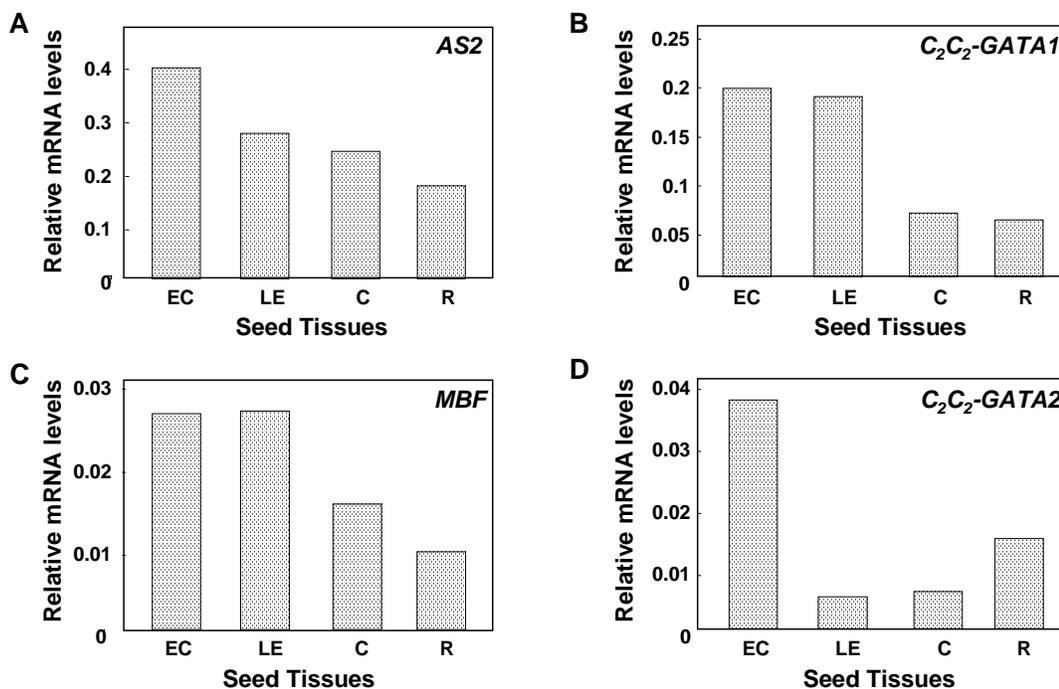


Figure 5. Tissue-specific expression of the four transcription factors in tomato seeds during germination. A, B, C and D, the expression of *GATA1*, *GATA2*, *MBF* and *AS2*, respectively. Expression was analyzed by quantitative real time PCR.

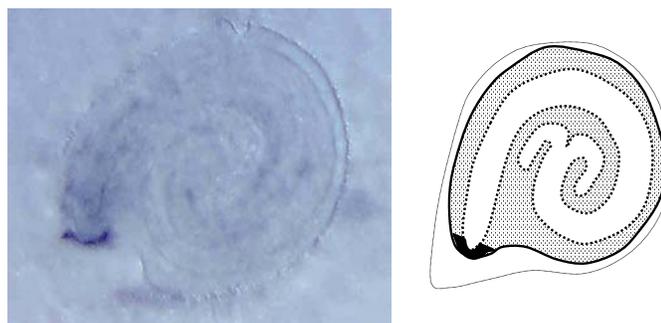


Figure 6. Tissue printing of *GATA2* expression. The expression was detected mainly in the endosperm cap and the radicle tip, which is consistent with the results of quantitative real time PCR in Fig. 5. Schematic representation of seed exhibiting micropylar specific expression is shown.

Bioinformatic analysis using the databases and sequence alignment revealed that the tomato *GATA2* zinc finger protein is similar to *AGP1*, a zinc finger protein in

tobacco. Alignment of the amino acid sequences of the tobacco NtAGP1 and tomato GATA2 (SIAGP1) showed an 84 % identity. The amino acid sequence of the zinc finger domain was conserved between the two proteins (Fig. 7).

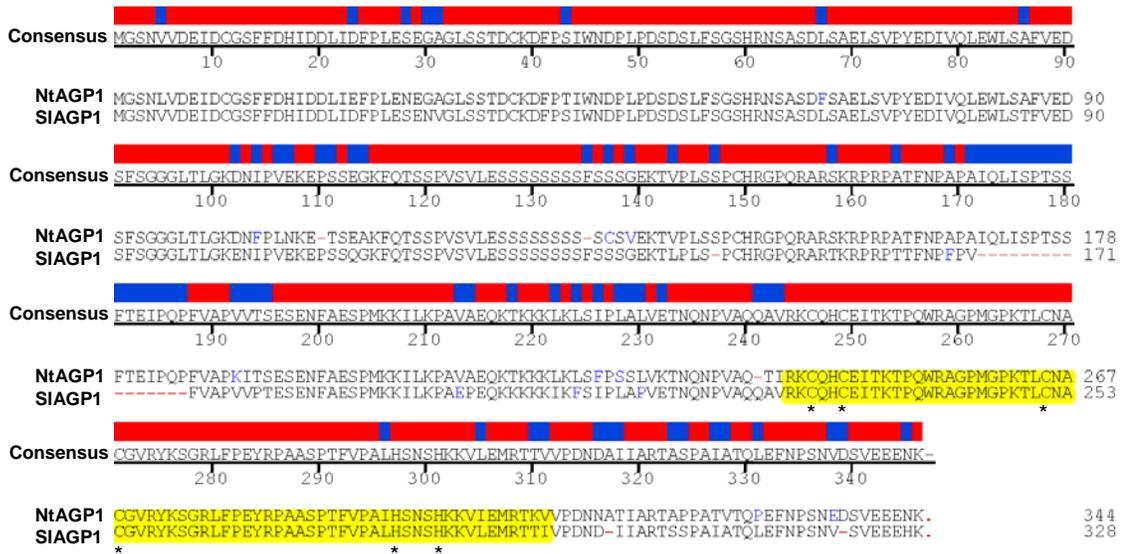


Figure 7. Alignment of the amino acid sequences of the tomato GATA2 (SIAGP1) and tobacco AGP1 (NtAGP1). Identical and distinct amino acids are highlighted in red and blue, respectively. The zinc finger domain is highlighted in yellow. Cys and His contributing to the zinc finger are marked by asterisks.

In tobacco, NtAGP1 binds the AG motif of the promoter region of *NtMyb2*, a Myb transcription factor (Sugimoto et al., 2003). An orthologue of *NtMyb2* in tomato was found in the tomato EST database (<http://compbio.dfc.harvard.edu/tgi/cgi-bin/tgi/gimain.pl?gudb=tomato>) (Fig. 8) and was expressed in tomato seeds (Fig. 9). Alignment of *SIMyb2* in tomato with *NtMyb2* in tobacco showed an 85% identity (Fig. 8).

```

SIMyb2      1      CACTG      CTA      TTTCT      AAAACACATC
NtMyb2     1      -----CACTA-----CTA----CTTCT----AAAGAACACC
1      atttcttgaagaaaagagagCACTgaaaaagagCTAaaac TTCTgttcaAAAA ACA C

SIMyb2     61      AAGAAACAGTACACT      AAGAATAAAACCAA      TAAAGAG      ATTT
NtMyb2    24      ---AATATTCAATTCATT----AAACATAAAAAAAA-----TAAAAA--ATTC-----
61      caaaa A CAgT CA TtctgAAg ATAAAA AAgtgatTAAAGAgtcATT gttcac

SIMyb2    121      ATCAAAGATT      TGATTAAA      AAAATGGT      GAGAGCTCCTTGTGTGAGAAGAT
NtMyb2    64      --ATCAAAGTT-----TTTTTAAA---AAAATGGTTAGAGCTCCTTGTGTGAGAAGAT
121     aaATCAAgaTTtagcaagT TTAAAgcaAAAATGGT AGAGCTCCTTGTGTGAGAAGAT

SIMyb2    181      GGGGTTGAAAAGGACCATGGACTCC      GAAGAAGATCAAATTCT      TGT      TCTTTA      TATTCA
NtMyb2    113     GGGGTTAAAGAAAGGACCATGGACTCC      GAAGAAGATCAAATTCT      CAT      CTCTTA      CATTCA
181     GGGGTTgAAaAAAGGaCCATGGACTCC GAAGAAGATCAAATTCT gT TCTTA ATTCA

SIMyb2    241      AACAAATGGCCATGGCAATTGGAGAGCCCTCCCTAAAC      TAGCTGGACTTTTGAGATGTGG
NtMyb2    173     ATCAAATGGTCATGGCAATTGGAGAGCCCTCCCTAAAT      TAGC      CGGATTA      TTGCGATGTGG
241     A CAAATGG CATGGCAATTGGAGAGCCCTCCCTAAA TAGC GGA T TTG GATGTGG

SIMyb2    301      GAAGAGTTGCAGATTGC      GTGGACTAAC      TATTTG      CGTCCAGATATA      AAAGAGGGGAAAC      TTT
NtMyb2    233     AAAAGAGTTGCAGACTTAGATGGACGAATTATTTA      CGTCCAGATATA      TAAAGAGGAAAT      TTT
301     gAAgAGTTGCAGA T G TGGAC AA TATTTgCGTCCAGATAT AAgAGgGGA AA TT

SIMyb2    361      TACTAGAGAAGAAGAAGACTCC      ATTATTCAG      TTACATGAAATGCTTGGCAATAGATGGTC
NtMyb2    293     CACAAGAGAAGAAGAAGATAGTATTATTCAA      TTACATGAAATGCTTGGCAATAGATGGTC
361     AC AGAGAAGAAGAAGA ATTATTCAGTTACATGAAATGCTTGGCAATAGATGGTC

SIMyb2    421      TGCAATAGCAGCGAGATTACCGGGACGAAC      GGA      CAATGAAAT      TAAAAATGTATGGCACAC
NtMyb2    353     TGCATAGCAGCTAGATTACCGGGACGAAC      AGATA      AATGAAAT      TAAAAATGTATGGCACAC
421     TGCaATAGCAGC AGATTACCGGGACGAACgGA AATGAAAT TAAAAATGTATGGCACAC

SIMyb2    481      CCACTTGAAAAAGAGGCTTAAAAATTA      C      CAGCCTCCTCAAAA      CTCCAAAAGACACTC      CAA
NtMyb2    413     TCACTTGAAAAAAGGCTTAAAAATTA      T      CAGCCTCCTCAAAAG      CTCCAAAAGACACTC      CAAA
481     CACTTGAAAAAagAGGCTTAAAAATTA CAGCCTCCTCAAAAaCTCCAAAAGACACTC AA

SIMyb2    541      AAACAAC      GATTCCAAAGCTCCT      AGTACTTCTCAAAA      ---      CCTTCAATA      ATTCAGACAA
NtMyb2    473     AAACAAG---GATTCCAAAGCTCCT      GTACTTCTCAAAA      CCTTGAAAAGTTCAAACAA
541     AAACAA cttGATTCCAAAGCTCCT GTACTTCTCAAAAttgCCTT AA AaTTCaGACAA

SIMyb2    598      TTTTAGCAATATC      CAAGAAGATATTAATGGGC      CCGTGA      CCGGCCCGAACTCGCCACAA      CG
NtMyb2    530     TTTTAGCAA      CATCAAGAAGA      CCGGCC      GGGCTTGGGT      CCGGCCCGAACTCGCCACAA      TT
601     TTTTAGCAA ATC AAGAAGA a GGGC G G CCGGCCCGAACTCGCCACAA

SIMyb2    658      ATCGTCTAGT      GAGATGTCGACTGTCACGG      TT      GATTCAAC      AGCCATGAC      AAT
NtMyb2    590     GTCATCTAGCGAGATGTCGACTGTCACGG      CC      GATTCA      CTAGCCGTGAC      -----      AAT
661     aTCgTCTAG GAGATGTCGACTGTCACGG GATTCA AGCCaTGACgaccatcacAAT

SIMyb2    718      CGATGATCAGAATA      GCAAT      GATGAGATGGACTC      G      TCTGAAAAATTTTATTCC
NtMyb2    641     GGCATCTCGAACA-----GTAAC--GACCAAAATAGACTCATCTGAAAAATTTTATTCC
721     GA g GAA Atgtttaag AA taGA AgATgGACTCgTCTGAAAAATTTTATTCC

```

(sequence continues to the next page)

Figure 8. Alignment of the nucleotide sequences of the tomato *Myb2* (*SIMyb2*) and tobacco *NtMyb2*. Conserved nucleotides are highlighted in black.

(sequence continued from previous page)

```

SIMyb2 778 AGAGATGATGAGAGTTTGGACGGATGATTTATCCACAAGCGATGG
NtMyb2 692 AGAGATCGATGAGAGTTCTGGACGGACGGTTTGTCCACGAGTGGT-----GG
781 AGAGAT GATGAGAGTTT TGGACGGA GaTTTaTCCACaAG GaTaaactcgacttttGG

SIMyb2 838 TGGTGAGAATTACAAGTCCAATTTCCATTTCTTCGGTGAAGCAAGA
NtMyb2 740 T-----GGTGAAGAATTACAAGTCCAATTTCCATTTCAATGACATGAAACAAGA
841 TatggaggggtaccGGTGgAGAATTACAAGT CAATTTCCATTT T gTGAAGCAAGA

SIMyb2 898 AAGTATGGACATGG--TTGGAGCAAATTAGAGGACGACATGGACTTTTGGTACAATGT
NtMyb2 788 AAATGTAGAGAAGGTTGGAGCAAATTAGAGGATGATATGGACTTTTGGTACAATGT
901 AAgTaTgGA A GGatgTTGGAGCAAATTAGAGGA GA ATGGACTTTTGGTACAATGT

SIMyb2 955 TTTCATAAAGTCGGGGACTTAC TAGATTTACCgGAATTTTGAGTGGTCAATTTGATTGT
NtMyb2 848 TTTCATAAAGTCGGGGACTTATTAGAGTTACCAgAAATTTTGAGGGTTAATTTGTTGT
961 TTTCATAAAGTC GGGGACTTA TAGA TTACCgGAATTTTGAG GGT AATTTGATTGT

SIMyb2 1015 A TACAAAACCTTGAAGTAGTGGAAATGCCAGCTAATTA-----
NtMyb2 908 T-TTCAAAAACCTTGAAGTAGTGGAAATGCCAACTAATTT
1021 aT CAAAACCTTGAAGTAGTGGAAATGCCAgCTAATT agtgggtgttttttttgggattt

SIMyb2 1052 -TTGGGAGTCAACAAGTTTGAAACTTCAATGTTTGTATTGACCTTTAC-CTCTTGATAG
NtMyb2 967 TTGGGAGTCGACAAGTTTGAAAATTTTGTTTGTTATTGACTTTAAA TCCTTGAGAG
1081 tTTGGGAGTCaACAAGTTTGAAA TT TGTTTGTATTGAC TT A g TCTTGA AG

SIMyb2 1110 GACCACCAAATAC----TACAAGTTGATACCTTCTTTTTTTTA-GTTAGGATAAT
NtMyb2 1027 GACCAACAAAAA TACAAGTTGATCCTTTTTTTTTTTT ACTAGGATAAT----
1141 GACCA CAAA A aaaaTACAAGTTGAT CTTT TTTTTTTT tg TAGGATAATctttt

SIMyb2 1165 TTTTTCTGTCTTTATTTT-AAACCTTTTAGTT-AG
NtMyb2 1082 -----TTTTCTATTTTTATTTT AAACCTTTTAGTT AG
1201 tcttttcttttcttttctactTTTTTCTGT TTTATTTTgAA C TTTTAGTTAG

SIMyb2 1223 TTTAATTGGGAGAAA GCATATAGTGGATGGTGATATGAAAAAAGAAA
NtMyb2 1119 TTTAATTGGGAGAAAAGTATAGTGGATGGTGATGTGAAAAAGGGGA-----
1261 TTTAATTGGGAgAAAg aTATAGTGGATGGTGATaTGAAAAAGaaAgattatgatggaa

SIMyb2 1283 AATTATTAGTAATATTA GGAAAAAAAGATTT
NtMyb2 1166 -AATTATTGGAGATATTG-----GGAAAGAAAAAATTT-----
1321 tAATTATTaG aATATTaattaggattaGGAAAAAAAGa TTagagaaaagacttcaa

SIMyb2 1343 ACTC
NtMyb2 1198 ---ATC-----
1381 gaaA Tctagtcaacatcctcctaacttagcttaattgtatgtgaattacctcttttgta

SIMyb2 1403
NtMyb2 -----
1441 acatgacattacccaaaaagaataaaaaatgttttcatttaaaaaaaaaaaaaaaaaaa

```

Figure 8. Alignment of the nucleotide sequences of the tomato *Myb2* (*SIMyb2*) and tobacco *NtMyb2*. Conserved nucleotides are highlighted in black.

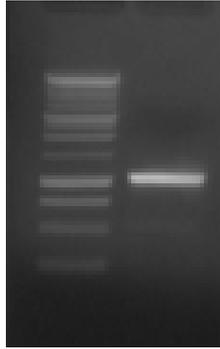


Figure 9. *SIMyb2* expression in germinating tomato seeds. Total RNA was extracted from 24 h-imbibed seeds and used for reverse transcription to synthesize cDNA. *SIMyb2* was amplified using Myb specific primers.

DISCUSSION

It is known that *LeMAN2* encodes for the endosperm cell wall modifying enzyme endo- β -mannanase and is induced by GA (Nonogaki et al., 2000). However, upstream regulators of *LeMAN2* are unknown. Taking advantage of the sequenced genome of *Arabidopsis thaliana*, GA-inducible transcription factors potentially associated with seed germination have been identified (Ogawa et al., 2003). Tomato orthologues of the known *Arabidopsis* GA-inducible transcription factors were found using bioinformatics. Quantitative real time PCR revealed that four transcription factors, *AS2*, *MBF*, *GATA1* and *GATA2* were expressed in tomato seeds during germination. Three of these transcription factors, *AS2*, *MBF* and *GATA1* were expressed mainly in the early stage of seed germination (12 h). In contrast, *GATA2*, a GATA zinc finger gene was highly expressed during the later stage of seed germination (36 h). The expression pattern of *GATA2* is very similar to that of *LeMAN2* whose expression reached the peak during the later stages of germination. Real time PCR using RNA extracted from the endosperm cap (EC), lateral endosperm (LC), radicle half embryo (R) and cotyledon half embryo (C) indicated that, while *GATA1*, *AS2* and *MBF* were expressed in all tissues to a certain

level, *GATA2* was expressed primarily in the endosperm cap (Fig. 5 and 6). Thus, the tissue specificity of the *GATA2* was also similar to that of *LeMAN2*. This was confirmed by tissue print analysis. These results suggest that *GATA2* could be involved in the regulation of *LeMAN2* expression.

The tomato *GATA2* transcription factor is a GATA zinc finger protein and similar to *AGP1*, an AG-motif binding protein in tobacco. The amino acid sequence of *GATA2* displayed an 84% identity with the amino acid sequence of *NtAGP1*. The zinc finger domain containing the characteristic 2 cysteines and 2 histidines was completely conserved between these two proteins (Fig. 7). It is possible that *GATA2* directly induces *LeMAN2* since a GATA motif is found in the promoter region of *LeMAN2* (Fig. 10). Alternatively, *GATA2* might be involved in the regulation of *LeMAN2* expression in an indirect manner through the activation of a direct regulator of *LeMAN2*.

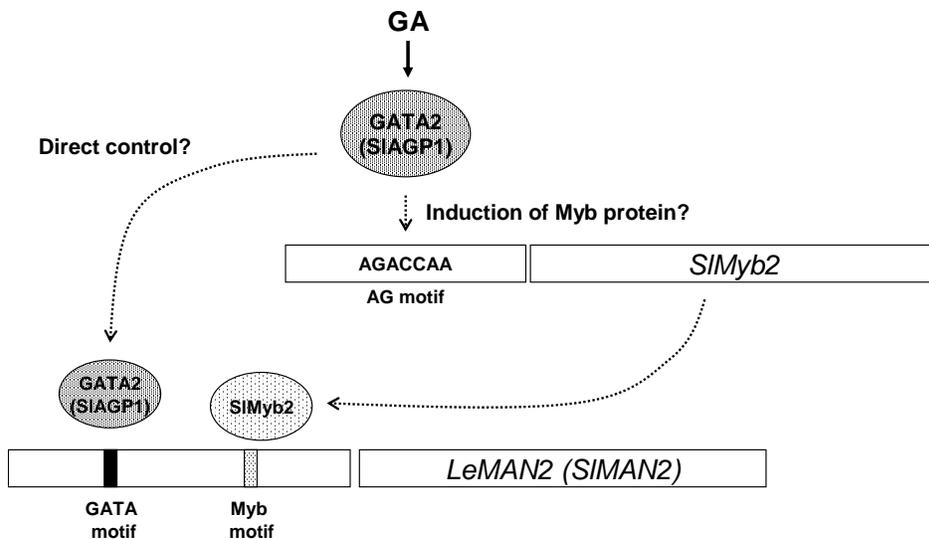


Figure 10. The hypothetical schematic of *LeMAN2* regulation by the transcription factor *GATA2*. *GATA2* could interact with the GATA motif in the promoter region of *LeMAN2* and directly activate *LeMAN2*. Alternatively, *GATA2* might be involved in the induction of *LeMAN2* in an indirect manner through the activation of a Myb protein.

In tobacco, NtAGP1, an orthologue of tomato GATA2, functions by binding to the AG motif (AGATCCAA) in the promoter region of *NtMyb2*, a *Myb* gene (Sugimoto et al., 2003). EST database search and RT PCR have shown that the tomato *Myb* gene, *SIMyb2* which shows an 85% identity to the tobacco *Myb* gene, is expressed during seed germination (Fig. 9). Therefore, it is possible that GATA2, a tomato orthologue of NtAGP1, binds to the promoter region of *SIMyb2* and activates this tomato *Myb* gene. *SIMyb2* might in turn bind to a *Myb* motif in the *LeMAN2* promoter and activate its expression (Fig 10). This hypothesis needs to be examined by testing whether the tomato GATA2 (SIAGP1) physically interacts with the AG motif-containing sequence in the promoter region of *SIMyb2*, and also whether *SIMyb2* could bind to the *Myb* motif-containing region of the *LeMAN2* promoter, with a gel mobility shift assay in future experiments.

To determine additional upstream regulators of *LeMAN2* and other genes involved in tomato seed germination, a tomato GeneChip with about 9,000 gene probes can be used. We are currently conducting hybridization of RNA extracted from the endosperm cap, lateral endosperm, radicle half embryo and cotyledon half embryo with the tomato GeneChip in collaboration with RIKEN Plant Science Center, Japan. From this analysis, additional genes expressed primarily in the endosperm cap will be identified. Future work will involve the characterization of genes identified from this analysis to elucidate the molecular mechanisms of seed germination.

Chapter 3

General Conclusion

Jing Sun

Germination is strictly defined as the physiological events prior to the emergence of the radicle. Radicle emergence occurs through the endosperm cap, the region of endosperm adjacent to the radicle tip. Seed germination occurs only when the growth potential of the embryo is greater than the opposing mechanical resistance of the endosperm cap. Understanding endosperm cell wall weakening is essential for determining the mechanisms involved in seed germination. In tomato endosperm, the major cell wall polysaccharide is galactomannan. Therefore, understanding the function and the regulation of *LeMAN2* endo- β -mannanase is critical to elucidate the mechanisms of seed germination in this species. While it is clear that *LeMAN2* is inducible by GA, the upstream regulators of *LeMAN2* and molecular mechanisms through which GA activates this gene are not understood. In this thesis research, potential regulators of *LeMAN2* have been identified and characterized. *GATA2*, a GATA zinc finger gene is expressed in a temporal and spatial manner similar to that of *LeMAN2* and is a potential upstream regulator of *LeMAN2*. The same GATA zinc finger domain is found in both tomato and tobacco AGP1. NtAGP1 binds to the AG motif in the promoter of *NtMyb2*. A similar Myb protein is expressed in germinating tomato seed. A Myb binding motif is found in the promoter region of *LeMAN2*. These new findings have provided new hypotheses concerning the potential direct and indirect regulation of *LeMAN2* during seed germination.

In addition, this thesis research also provided a good example of translational biology where the information obtained from the model plant *Arabidopsis* was translated into tomato, a model crop species. This type of approach can be expanded to other crop

species which are important for agriculture and our diet and also wild species which are important for our environment.

Bibliography

- Bewley JD** (1997) Breaking down the walls - a role for endo- β -mannanase in release from seed dormancy? *Trends in Plant Science* **2**: 464-469.
- Bewley JD, Black M** (1994) Mobilization of Stored Seed Reserves. *In*: J.D. Bewley and M. Black, eds, *Seeds : Physiology of development and germination*. Plenum Press, New York. pp. 293-343.
- Cantliffe DJ, Fischer JM, Nell TA** (1984) Mechanism of seed priming in circumventing thermodormancy in lettuce. *Plant Physiology* **75**: 290-294.
- Chen F, Bradford KJ** (2000) Expression of an expansin is associated with endosperm weakening during tomato seed germination. *Plant Physiology* **124**: 1265-1274.
- Crozier A, Kamiya Y, Bishop G, Yokota T** (2000) Biosynthesis of hormones and elicitor molecules. *In*: B.B. Buchanan, W. Gruissem and R.L. Jones, eds, *Biochemistry and Molecular Biology of Plants*. American Society of Plant Biologists, Rockville. pp.850-929.
- Debeaujon I, Koornneef M** (2000) Gibberellin requirement for Arabidopsis seed germination is determined both by testa characteristics and embryonic abscisic acid. *Plant Physiology* **122**: 415-424.
- Debeaujon I, Lepiniec L, Pourcel L, Routaboul JM**, eds (2007) Seed coat development and dormancy. *In*: K.J. Bradford and H. Nonogaki, eds, *Seed Development, Dormancy and Germination*. Blackwell Publishing, Oxford. pp. 25-49.
- Grant G, More LJ, McKenzie NH, Steward J** (1983) A survey of the nutritional and haemagglutination properties of legume seeds generally available in the UK. *British Journal of Nutrition* **50**: 207-214.

- Groot SPC, Karssen CM** (1987) Gibberellins regulate seed germination in tomato by endosperm weakening: a study with gibberellin-deficient mutants. *Planta* **171**: 525-531.
- Groot SPC, Kieliszewska-Rokicka B, Vermeer E, Karssen CM** (1988) Gibberellin-induced hydrolysis of endosperm cell walls in gibberellin-deficient tomato seeds prior to radicle protrusion. *Planta* **174**: 500-504.
- Nonogaki H, Bradford KJ** (2003) Tissue printing for localization of mRNA expression in seeds. In K.J. Bradford, D. Come, G. Nicolas, H. Pritchard, eds, *The Biology of Seeds: Recent Research Advances*. CAB International, Wallingford, pp. 171-179.
- Nonogaki H, Chen F, Bradford KJ** (2007) Mechanisms and genes involved in germination *sensu stricto*. In: K.J. Bradford and H. Nonogaki, eds, *Seed Development, Dormancy and Germination*. Blackwell Publishing, Oxford. pp. 254-304.
- Nonogaki H, Gee OH, Bradford KJ** (2000) A Germination-specific endo- β -mannanase gene is expressed in the micropylar endosperm cap of tomato seeds. *Plant Physiology* **123**: 1235-1246.
- Ogawa M, Hanada A, Yamauchi Y, Kuwahara A, Kamiya Y, Yamaguchi S** (2003) Gibberellin biosynthesis and response during Arabidopsis seed germination. *The Plant Cell* **15**: 1591-1604.
- Sambrook J, Fritsch EF, Maniatis T** (1989) *Molecular cloning: A Laboratory Manual* (2nd edition). Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Sanchez RA, Sunell L, Labavitch JM, Bonner BA (1990) Changes in the endosperm cell walls of two *Datura* species before radicle protrusion. *Plant Physiology* **93**: 89-97.

Sugimoto K, Takeda S, Hirochika H (2003) Transcriptional activation mediated by binding of a plant GATA-type zinc finger protein AGP1 to the AG-motif (AGATCCAA) of the wound-inducible Myb gene NtMyb2. *The Plant Journal*. **36**: 550-564.

Watkins JT, Cantliffe DJ, Huber DJ, Nell TA (1985) Gibberellic acid stimulated degradation of endosperm in pepper. *Journal of the American Society for Horticultural Science* **110**: 61-65.

Yamaguchi S (2008) Gibberellin metabolism and its regulation. *Annual Review of Plant Biology* **59**: 225-251.